D1-82-0159

BOEING SCIENTIFIC RESEARCH LABORATORIES

Enclosure Theorems for Eigenvalues and Probability Inequalities

Albert W. Marshall

Ingram Olkin

Mathematics Research

April 1962

D1-82-0159

ENCLOSURE THEOREMS FOR EIGENVALUES AND PROBABILITY INEQUALITIES*

by

Albert W. Marshall

and

Ingram Olkin
Stanford University

Mathematical Note No. 255

Mathematics Research Laboratory

BOEING SCIENTIFIC RESEARCH LABORATORIES

April 1962

*This report is also being issued as Technical Report No. 1, prepared under the auspices of National Science Foundation Grant 19138, Applied Mathematics and Statistics Laboratories, Stanford University, Stanford, California, May 7, 1962.

ENCLOSURE THEOREMS FOR EIGENVALUES

AND PROBABILITY INEQUALITIES

by

Albert W. Marshall and Ingram Olkin 1/

Boeing Scientific Research Laboratories

and

Stanford University

1. Introduction.

Let H be an Hermitian matrix, z a row vector of unit length, $\mu = zHz^*$, $\sigma^2 = z(H - \mu I)^2 z^*$. The enclosure theorem of Krylov and Bogoliubov [1] (see also Weinstein [2], Kohn [3], Block and Fuchs [4], Swanson [5]) asserts that there is an eigenvalue in the interval $[\mu - \sigma , \mu + \sigma]$.

We now present a probabilistic interpretation which yields a simple proof of this result as well as other enclosure theorems.

Write $H = U^*DU$, where U is unitary, $D = \operatorname{diag}(\theta_1, \ldots , \theta_k)$, then for any unit vector z ,

$$\mu = zHz^* = (zU^*) D(Uz^*) \equiv \sum p_j \theta_j ,$$

with $0 \leq p_j$, $\sum p_j = 1$. Consequently, we may view μ as the mean of a random variable X with probability distribution $P\{X = \theta_j\} = p_j$, $j = 1, \ldots , k$. Similarly, $\sigma^2 = z(H - \mu I)^2 z^* = \sum p_j(\theta_j - \mu)^2$ is the variance of X . According to the Bienaymé-Chebyshev inequality,

1/ The work of the second author was supported in part by the National Science Foundation.

$$P(\,|X - \mu| \le \sigma t\,) \ge 1 - t^{-2} \quad . \tag{1}$$

Hence, for all $t > 1$ there is positive mass in the interval $[\mu - \sigma t\ ,\ \mu + \sigma t]$, which implies that there is positive mass in $[\mu - \sigma\ ,\ \mu + \sigma]$. Since the mass is distributed only on eigenvalues, this means that there is at least one eigenvalue in $[\mu - \sigma\ ,\ \mu + \sigma]$.

In general, if a bound is obtained for some set $\mathcal{S}$, i.e., $P(X \in \mathcal{S}) \ge p$, then $\sum_{\mathcal{S}} p_j \ge p$, where the summation is over all j corresponding to $\theta_j \in \mathcal{S}$. But $p_j = |u_j\, z^*|^2$, where u_j is the j^{th} row of U , and is, therefore, an eigenvector corresponding to the eigenvalue θ_j , so that $\sum_{\mathcal{S}} |u_j\, z^*|^2 \ge p$.

In Section 2 we state a number of enclosure theorems, and in Section 3 we present the related probabilistic inequalities which furnish proofs.

2. Enclosure Theorems.

We now state a number of enclosure theorems to illustrate the variety of results which may be obtained in this manner.

In the following we let H be an Hermitian matrix, z a row vector of unit length, $\mu = zHz^*$, $\sigma^2 = z(H - \mu I)^2 z^*$.

2.1 General Hermitian Matrix.

There exists at least one eigenvalue in the interval

$$\left[a + \mu - \sqrt{\sigma^2 + a^2}\ ,\ a + \mu + \sqrt{\sigma^2 + a^2}\right] .$$

Remark.

In this and the following theorems, a slightly stronger version can be obtained by making the interval half-open.

This theorem is given by Kato [6]. The result of Section 1 is a special case with $a = 0$.

2.2 Positive Definite Hermitian Matrix.

Let H be positive definite, and $\mu_{-1} = zH^{-1}z^*$. If $s < \mu$, $s < 1/\mu_{-1}$, then the interval $[s, (\mu - s) / (1 - s\mu_{-1})]$ contains at least one eigenvalue.

The particular interval, $[(1 - \sqrt{\mu\mu_{-1} - 1}) / \mu_{-1}, (1 + \sqrt{\mu\mu_{-1} - 1}) / \mu_{-1}]$, has minimum length.

2.3 Bounded Hermitian Matrix.

Let H and $I - H$ be positive definite. If $-\mu \leq r \leq -\sigma^2/(1-\mu)$, then there is at least one eigenvalue in $[0, r+\mu] \cup [-\sigma^2/r + \mu, 1]$.

3. Probability Inequalities.

We now present the probability inequalities which furnish proofs for the enclosure theorems of Section 2.

3.1 Theorem of Selberg [7].

If X is a random variable with $EX = \mu$, $E(X-\mu)^2 = \sigma^2$, then for all real a and positive b,

$$\text{(i)} \quad P(|X - \mu - a| < b) \geq 1 - \frac{\sigma^2 + a^2}{b^2}, \quad \text{if} \quad \sigma^2 \geq |a|\, b - a^2,$$

$$\text{(ii)} \quad P(|X - \mu - a| < b) \geq 1 - \frac{\sigma^2}{\sigma^2 + (b - |a|)^2}, \quad \text{if} \quad \sigma^2 < |a|\, b - a^2.$$

This inequality is sharp.

The proof of this asymmetric version of the Bienaymé-Chebyshev inequality in [7] is unnecessarily complicated, and we give the following proof.

<u>Proof</u>.

If $f(x) = (x-a)^2 / b^2$ whenever $\sigma^2 \geq |a| \; b - a^2$, $f(x) = [(b - |a|) \; x - \sigma^2]^2 / [(b - |a|)^2 + \sigma^2]^2$ whenever $\sigma^2 < |a| \; b - a^2$, then $f(x) \geq 0$ for all x and $f(x) \geq 1$ for all $|x - \mu - a| \geq b$. The result follows from $E \; f(X) \geq P(\; |X - \mu - a| \geq b \;)$.

If $b^2 - a^2 \geq \sigma^2 \geq |a| \; b - a^2$, the extremal distribution is $P(\; Y - \mu = a + b \;) = p_1$, $P(\; Y - \mu = a - b \;) = p_2$, $P(\; Y - \mu = a \;) = 1 - p_1 - p_2$, where $p_1 = (\sigma^2 - ab + a^2) \; / \; 2b^2$, $p_2 = (\sigma^2 + ab + a^2) / 2b^2$.

If $\sigma^2 < |a| \; b - a^2$, $a > 0$, the extremal distribution is $P(\; Y - \mu = a - b \;) = \sigma^2[\; (b-a)^2 + \sigma^2 \;]^{-1} = 1 - P(\; Y - \mu = \sigma^2 \; / \; (b-a) \;)$. For $a < 0$, use $-Y$ for Y.

If $\sigma^2 > b^2 - a^2$, the extremal distribution is $P(\; Y - \mu = a + b - c \;) = (b + c - a) \; / \; [\; 2(b+c) \;] = 1 - P(\; Y - \mu = a - b - c \;)$, where $c = \sqrt{\sigma^2 + a^2} - b$. ‖

To obtain the enclosure theorem, we consider (i) and require $(\sigma^2 + a^2) \; / \; b^2 < 1$, so that for all $b > \sqrt{\sigma^2 + a^2}$, there is an eigenvalue in the interval $[\; a + \mu - \sqrt{\sigma^2 + a^2} \; , \; a + \mu + \sqrt{\sigma^2 + a^2} \;]$.

Since the bound (ii) is always less than 1, we obtain the shortest interval by choosing $b = (\sigma^2 + a^2) \; / \; |a|$. But this leads to the interval $[\; a + \mu - (\sigma^2 + a^2) \; / \; |a| \; , \; a + \mu + (\sigma^2 + a^2) \; / \; |a| \;]$ which is contained in $[\; a + \mu - \sqrt{\sigma^2 + a^2} \; , \; a + \mu + \sqrt{\sigma^2 + a^2} \;]$.

The extremal distribution for (i) is unique and places positive probability at both of the points $a + \mu - b$ and $a + \mu + b$. If we eliminate this extremal distribution, the probability is strictly less than the bound, so that $P[\ |X - \mu - a| < \sqrt{\sigma^2 + a^2}\ ] > 0$, which says that there is an eigenvalue in the open interval $(\ a + \mu - \sqrt{\sigma^2 + a^2}\ ,\ a + \mu + \sqrt{\sigma^2 + a^2}\ )$. By using a half-open interval, the extremal distribution is included.

3.2 <u>An Inequality with Negative Moments</u>.

If X is a positive random variable, i.e., $P\{X \leq 0\} = 0$, $EX = \mu$, $EX^{-1} = \mu_{-1}$, then

$$P[\ s < X < t\ ] \geq 1 - \frac{\mu - 2\sqrt{st} + st\mu_{-1}}{(\ \sqrt{t} - \sqrt{s}\ )^2} .$$

Equality is achieved if and only if the bound is non-negative and X has the distribution $P\{X = s\} = (a + bt)\sqrt{s} / (t - s)$, $P\{X = t\} = (a + bs)\sqrt{t} / (t - s)$, $P\{X = \sqrt{st}\} = 1 - (a + b\sqrt{st}) / (\sqrt{t} - \sqrt{s})$, where $a = (\mu - \sqrt{st}) / (\sqrt{t} - \sqrt{s})$, $b = (\mu_{-1}\sqrt{st} - 1) / (\sqrt{t} - \sqrt{s})$.

<u>Proof</u>.

If $f(x) = (x - \sqrt{st})^2 / [x(\sqrt{t} - \sqrt{s})^2\ ]$, then $f(x) \geq 0$ for all $x > 0$, $f(x) \geq 1$ for all $x \notin (s\ ,\ t)$, and the result follows from $E\ f(X) \geq P[\ X \notin (s\ ,\ t)]$. The extremal distribution is then determined by placing probability where $f(x)$ is zero or one.

We note that $\mu\mu_{-1} \geq 1$, by the Schwarz inequality, so that $s < \mu < t$, $s < 1/\mu_{-1} < t$, follow from the condition that the bound be non-negative. ‖

To obtain the corresponding enclosure theorem, note that the bound is non-negative when $t \geq (\mu - s) / (1 - \mu_{-1}s)$.

3.3 Bounded Random Variables.

If X is a bounded random variable, i.e., $P(\alpha \leq X - \mu \leq \beta) = 1$, $EX = \mu$, $E(x - \mu)^2 = \sigma^2$, $\alpha \leq s \leq t \leq \beta$, then $P(X < s \text{ or } X > t) \geq p$ given as follows:

(i) $1 + \dfrac{\alpha\beta + \sigma^2}{(s-\alpha)(\beta - s)}$, if $s + t \leq \alpha + \beta$, $\sigma^2 \geq -\alpha s$, $\sigma^2 \geq -\beta s$,

(ii) $1 + \dfrac{\alpha\beta + \sigma^2}{(t-\alpha)(\beta - t)}$, if $s + t \geq \alpha + \beta$, $\sigma^2 \geq -\alpha t$, $\sigma^2 \geq -\beta t$,

(iii) $\dfrac{st + \sigma^2}{(\beta - s)(\beta - t)}$, if $s + t \leq \alpha + \beta$, $-\beta t \leq \sigma^2 \leq -\beta s$, $\sigma^2 \geq -st$,

(iv) $\dfrac{st + \sigma^2}{(s-\alpha)(t-\alpha)}$, if $s \quad t \geq \alpha + \beta$, $-\alpha s \leq \sigma^2 \leq -\alpha t$, $\sigma^2 \geq -st$,

(v) $\dfrac{s^2}{s^2 + \sigma^2}$, if $\sigma^2 < -\alpha s$,

(vi) $\dfrac{t^2}{t^2 + \sigma^2}$, if $\sigma^2 < -\beta t$,

(vii) 0, otherwise.

This inequality is sharp.

Proof.

Write $y = x - \mu$ and let

(i) $f(y) = (y - \alpha)(y - \beta) / [(s - \alpha)(\beta - s)]$,

(ii) $f(y) = -(y - \alpha)(y - \beta) / [(\beta - t)(t - \alpha)]$,

(iii) $f(y) = 1 - (y - s)(y - t) / [(\beta - s)(\beta - t)]$,

(iv) $f(y) = 1 - (y - s)(y - t) / [(t - \alpha)(s - \alpha)]$,

(v) $f(y) = (\alpha y + \sigma^2)^2 / (s^2 + \sigma^2)^2$,

(vi) $f(y) = (ty + \sigma^2)^2 / (t^2 + \sigma^2)^2$, then, for each respective case, $f(y) \geq 0$ for all $x \in [\alpha, \beta]$, and $f(y) \geq 1$ for all $x \in [s, t]$, so that $E f(X) \geq P(X \notin [s, t])$.

Sharpness is exhibited by the following examples. Let Z be a random variable with the distribution

$$P\{Z - \mu = a_1\} = (\sigma^2 + a_2 a_3) / [(a_3 - a_1)(a_2 - a_1)] ,$$

$$P\{Z - \mu = a_2\} = (\sigma^2 + a_1 a_3) / [(a_2 - a_3)(a_2 - a_1)] ,$$

$$P\{Z - \mu = a_3\} = (\sigma^2 + a_1 a_2) / [(a_3 - a_2)(a_3 - a_1)] .$$

The hypotheses are satisfied and equality is attained in cases (i) - (iv) if (a_1, a_2, a_3) is taken to be: (α, t, β) for (i), (s, t, β) for (ii), (β, t, α) for (iii), (t, s, α) for (iv).

Now let $P\{Z - \mu = -\sigma^2/a\} = p$, $P\{Z - \mu = a\} = 1 - p$. With $a = s$ and $a = t$, equality is attained in cases (v) and (vi), respectively. Equality in the case (vii) is attained if $P\{Z - \mu = s\} = -\sigma^2 / [s(t - s)]$, $P\{Z - \mu = t\} = \sigma^2 / [t(t - s)]$.

It should be noted that $\sigma^2 \leq -\alpha\beta$ since $\alpha \leq X - \mu \leq \beta$. ||

(This inequality, as well as others, has been obtained by Melvin Dresher and Albert Madansky using the theory of moment spaces.)

To obtain the corresponding enclosure Theorem 2.3, we examine the constraints which guarantee that the bound is non-negative. The intervals obtained for cases (i), (ii), (v) and (vi) are larger than those for cases (iii) and (iv). The latter two combined yield the result that there is an eigenvalue of $H - \mu I$ in $[\alpha, r] \cup [-\sigma^2 / r, \beta]$ for $\alpha \leq r \leq -\sigma^2/\beta$, i.e., there is an eigenvalue of H in

$[\alpha + \mu , r + \mu] \bigcup [\mu - \sigma^2 / r , \beta + \mu]$. The hypotheses $\alpha + \mu = 0$, $\beta + \mu = 1$ give the desired result.

4. An Enclosure Theorem for Pairs of Eigenvalues.

The following theorem is due to Hájek and Rényi [8].

4.1 Theorem.

If $X_1 , \cdots , X_n$ are independent random variables, $EX_j = 0$, $EX_j^2 = \sigma_j^2$, $0 \leq t_1\sigma_1 \leq \cdots \leq t_n\sigma_n$, then

$$P(|X_1 + \cdots + X_m| \leq t_m\sigma_m , m = 1 , 2 , \ldots , n) \geq 1 - \sum_1^n t_j^{-2} .$$

To obtain the corresponding enclosure theorem, define $\mu_i = x_i H x_i^*$, $\sigma_i^2 = x_i(H - \mu_i I)^2 x_i^*$, $i = 1 , 2$. Hence there is a pair of eigenvalues, not necessarily distinct, in the region $\mu_1 \pm t_1\sigma_1 , \mu_1 + \mu_2 \pm t_2\sigma_2$, for all $0 \leq t_1\sigma_1 \leq t_2\sigma_2 , t_1^{-2} + t_2^{-2} \leq 1$.

Thus, for example, if $t_1^2 = 1/\lambda$, $t_2^2 = 1 / (1 - \lambda)$, with $\sigma_1^2 / \sigma_2^2 \leq \lambda / (1 - \lambda)$, we obtain the region

$$\mu_1 \pm \sigma_1 / \sqrt{\lambda} , \mu_1 + \mu_2 \pm \sigma_2 / \sqrt{1 - \lambda} .$$

5. Remarks on Kantorovich's Inequality.

We now show how the probabilistic interpretation yields results concerning inequalities for quadratic forms.

If H is an Hermitian matrix such that $0 < m \leq zHz^* \leq M$ when z is a vector of unit length, Kantorovich's inequality [9] asserts

$$zHz^* \, zH^{-1}z^* \le (m+M)^2/(4mM)$$

The probabilistic interpretation of this is that if X is a random variable with $P\{m \le X \le M\} = 1$ $(m > 0)$, then

$$\text{(5.1)} \qquad EX \, EX^{-1} \le (m+M)^2 / (4mM) .$$

This inequality cannot be improved. However, if $EX = \mu$ is known, then the inequality from (5.1)

$$EX^{-1} \le (m+M)^2/(4mM\mu)$$

is not sharp, but there is an improvement, namely,

$$\text{(5.2)} \qquad EX^{-1} \le (m+M-\mu) / (mM) .$$

This inequality is based upon the simple fact that if $f(x) \ge g(x)$ for $x \in a$, then $\int_a f(x)\, d\mu \ge \int_a g(x)\, d\mu$. In particular, if $g(x)$ is convex on the interval (m, M) and $f(x)$ is the chord through the points $(m, g(m))$ and $(M, g(M))$, we obtain

$$\text{(5.3)} \qquad E\, g(X) \le \frac{g(M) - g(m)}{M - m} EX + \frac{M\, g(m) - m\, g(M)}{M - m} .$$

Equality is attained for the distribution $P\{X = m\} = 1 - P\{X = M\} = (M - \mu) / (M - m)$.

Of interest are the special cases $g(x) = x^{-1}$ and $g(x) = x^k$, $k > 1$.

A number of proofs of (5.1) have been given, e.g., see Henrici [10]. We note that a simple proof is based upon (5.2), since

$$EX\, EX^{-1} \le \left(\frac{m+M}{mM}\right)\mu - \frac{\mu^2}{mM} \le \frac{(m+M)^2}{4mM} \; .$$

Another form for (5.1) is the following: If $P\,(m \le X \le M) = 1$, $(m > 0)$, $P\,(0 < Z) = 1$, $EZ = 1$, then

$$\text{(5.4)} \qquad E\, ZX\; E\, ZX^{-1} \le \frac{(m+M)^2}{4mM} \; .$$

The proof is essentially the same as for (5.1). A direct application of this inequality is the case where A and B are permutable Hermitian matrices with $0 < m_1 \le zAz^* \le M_1$, $0 < m_2 \le zBz^* \le M_2$, $zz^* = 1$, whence

$$zA^2z^*\; zB^2z^* \le \frac{(m_1m_2 + M_1M_2)^2}{4m_1m_2M_1M_2}\,(zABz^*)^2 \; .$$

REFERENCES

[1] N. Krylov and N. Bogoliubov, "Sur le calcul des racines de la transcendante de Fredholm les plus voisines d'une nombre donné par les méthodes des moindres carrés et de l'algorithme variationel," Izv. Akad. Nauk SSSR, Leningrad (1929), pp. 471-488.

[2] D. H. Weinstein, "Modified Ritz method," Proc. Nat. Acad. Sci. U. S. A., 20 (1934), pp. 529-532.

[3] W. Kohn, "A note on Weinstein's variational method," Phys. Rev. 71 (1947), pp. 902-904.

[4] H. D. Block and W. H. J. Fuchs, "An enclosure theorem for eigenvalues," Bull. Amer. Math. Soc., 67 (1961), pp. 425-426.

[5] C. A. Swanson, "An inequality for linear transformations with eigenvalues," Bull. Amer. Math. Soc., 67 (1961), pp. 607-608.

[6] T. Kato, "On the upper and lower bounds for eigenvalues," J. Phys. Soc. Japan, 4 (1949), pp. 415-438.

[7] H. L. Selberg, "Zwei Ungleichungen zur ergänzung des Tchebycheffschen lemmas," Skand. Aktuarietidskr., 23 (1940), pp. 121-125.

[8] J. Hájek and A. Rényi, "Generalization of an inequality of Kolmogorov," Acta Math. Acad. Sci. Hungar., 6 (1955), pp. 281-283.

[9] L. V. Kantorovich, "Functional analysis and applied mathematics," Uspehi Mat. Nauk., 3 (1948), p. 89.

[10] P. Henrici, "Two remarks on the Kantorovich inequality," Amer. Math. Monthly, 68(1961), pp. 904-906.